

# Sidon-type conditions on set systems

Peter Dukes and Jane Wodlinger\*  
 Mathematics and Statistics  
 University of Victoria  
 Victoria, BC V8W 3R4  
 Canada

2 October, 2012

## Abstract

Consider systems of  $k$  element subsets (or blocks) on a base set of  $v$  points. If all  $t$ -subsets occur with the same frequency, one obtains a  $t$ -design. On the other hand, following Sarvate and Beam, systems in which all  $t$ -subsets occur with different frequencies are  $t$ -*adesigns*.

Here, we observe that  $t$ -adesigns always exist and ask for the smallest possible maximum frequency  $\mu$ . Emphasis is placed on the asymptotic behavior (in  $v$ ) of  $\mu$ . Exact results are obtained for  $t = 1$  from basic principles. Nearly optimal results (up to a constant multiple) are obtained for  $t = 2$  using PBD closure. Weaker, yet still reasonable bounds for higher  $t$  follow from a linear algebraic argument.

Some connections are made with the famous Sidon problem of additive number theory.

## 1 Introduction

Given a set system  $(X, \mathcal{A})$ , in which  $X$  is a finite set and  $\mathcal{A}$  is a collection of subsets of  $X$ , the *frequency* of a set  $T \subset X$  is the number of elements of  $\mathcal{A}$ , counting multiplicity, which contain  $T$ .

A  $t$ -*design*, or  $S_\lambda(t, k, v)$ , is a pair  $(V, \mathcal{B})$ , where  $V$  is a  $v$ -set of *points*,  $\mathcal{B}$  is a collection of  $k$ -subsets of  $V$ , called *blocks*, and having the property

---

\*Research of the authors is supported by NSERC.

that every  $t$ -subset has frequency  $\lambda$ . Repeated blocks are permitted in this definition.

It is well-known and easy to see that an  $S_\lambda(t, k, v)$  exists only if

$$\lambda_i = \frac{\binom{v-i}{t-i}}{\binom{k-i}{t-i}}$$

is an integer for each  $i = 0, 1, \dots, t$ .

In [4], an interesting condition for set systems is considered. A  $t$ -*adesign* is a set system  $(V, \mathcal{A})$ , where as before  $V$  is a set of  $v$  points and  $\mathcal{A}$  is a collection of blocks of size  $k$ , having the condition that every  $t$ -subset of points has a **different** frequency.

Here, we abbreviate a  $t$ -adesign with  $A(t, k, v)$ . It is fairly easy to see that adesigns always exist.

**Proposition 1.1.** *An  $A(t, k, v)$  exists for all integers  $v \geq k \geq t \geq 1$ .*

*Proof.* List the  $k$ -subsets of  $V$  as  $K_0, K_1, \dots, K_{\binom{v}{k}-1}$ . Assign  $K_i$  multiplicity  $2^i$ . Clearly, each  $t$ -subset of points is contained in a different collection of  $k$ -subsets. So the frequencies

$$f(T) = \sum_{K_i \supseteq T} 2^i,$$

are distinct integers as  $T$  varies over  $\binom{X}{t}$ . □

This begs a more intricate question. Let  $\mu(t, k, v)$  denote the smallest maximum frequency, taken over all adesigns  $A(t, k, v)$ . The main question motivating this article is the following.

Given  $t, k, v$ , determine (or bound)  $\mu(t, k, v)$ .

In most of the previous investigations on adesigns, the cases of interest have been for  $t = 2$  and when the different pairwise frequencies are  $1, 2, \dots, \binom{v}{2}$ . It should be noted that here we allow zero as a frequency, although similar ideas hold for adesigns with nonzero frequencies.

An obvious lower bound is

$$\mu(t, k, v) \geq \binom{v}{t} - 1.$$

The proof of Proposition 1.1 shows that

$$\mu(t, k, v) < 2^{\binom{v}{k}};$$

however, we find this upper bound unsatisfactory. Not only is it non-polynomial in  $v$ , but it also fails to take advantage of the possibility of  $t \ll k$ .

In Section 2, we begin by explicitly calculating  $\mu(1, k, v)$  in all cases. This is mostly a review and summary of earlier work. We also show using PBD-closure techniques that  $\mu(2, k, v)$  is of order  $v^2$ , as expected. The outstanding work for  $t = 2$  essentially amounts to a reduction in the multiplicative constant, although this seems rather difficult with current techniques. Then in Section 3, we apply some linear algebra to obtain a weaker (yet still polynomial) bound for larger  $t$ .

Before beginning our detailed investigations, we should mention some connections with a beautiful and important topic in additive combinatorics. Briefly, a *Sidon set* is a set  $S$  of positive integers whose pairwise sums are all distinct, up to swapping summands. The “Sidon problem” of constructing Sidon sets with small elements is here extended to a combinatorial version. We wish to equip the  $k$ -subsets of a finite set with integer multiplicities so that the frequencies of all  $t$ -subsets (i.e. the inclusion sums) are all distinct.

In the number theoretic problem, it is known that the largest cardinality  $R(n)$  of a Sidon set  $S \subseteq \{1, \dots, n\}$  is roughly  $\sqrt{n}$ .

Now consider an adesign  $A(t, t+1, t+2)$ , where  $V$  is the ground set of size  $t+2$ . For  $t+2 \leq R(n)$ , assign weight  $f(x)$ , chosen from a Sidon set, to the co-singleton set  $\{x\}^c$ ,  $x \in V$ . Then it follows that the inherited weight on  $t$ -subset  $\{x, y\}^c$  is  $f(x) + f(y)$ . By construction, this takes distinct values on all  $t$ -subsets. It follows that  $\mu(t, t+1, t+2) \leq Ct^2$ . This is best possible, up to a constant multiple; however, we see that the exact determination of  $\mu(t, t+1, t+2)$  is as difficult as the Sidon problem.

Indeed, adesigns with the minimum possible  $v = k+1$  are related to  $(B_{v-t})$ -sequences; see [3] for more details on this and other variants of the number theoretic Sidon problem.

## 2 $t = 1$ and $t = 2$

Starting from a very simple construction due to Sarvate and Beam, it is possible to completely solve the maximum frequency problem for 1-adesigns.

The general strategy is as follows. Suppose  $f(1) < \dots < f(v)$  are prescribed frequencies whose sum  $F$  is divisible by  $k$ . Set up  $b = F/k$  blocks, and place element 1 in the first  $f(1)$  blocks, element 2 in the next  $f(2)$  blocks, and so on, with blocks identified modulo  $b$ . In other words, block  $i$  contains those elements  $x$  such that

$$\sum_{1 \leq y < x} f(y) < bq + i \leq \sum_{1 \leq y \leq x} f(y)$$

for some nonnegative integer  $q$ . Care must be taken that the maximum frequency  $m = f(v)$  does not exceed  $b$ , the number of blocks. Ideally, the frequencies are chosen to be consecutive, or almost consecutive integers.

**Theorem 2.1.**

$$\mu(1, k, v) = \begin{cases} v - 1 & \text{if } 2k \leq v \text{ and } \binom{v}{2} \equiv 0 \pmod{k}, \\ v & \text{if } 2k \leq v \text{ and } \binom{v}{2} \not\equiv 0 \pmod{k}, \\ \left\lceil \frac{1}{v-k} \binom{v}{2} \right\rceil & \text{if } 2k > v. \end{cases}$$

*Proof.* There is a natural division into two main cases.

CASE 1:  $2k \leq v$ . If  $k \mid \binom{v}{2}$ , then simply use the above strategy with  $b = \binom{v}{2}/k$  blocks and point frequencies  $0, 1, \dots, v - 1$ . Note that  $b \geq v - 1$  follows from the assumption  $2k \leq v$ . On the other hand, if  $\binom{v}{2} = bk - r$ ,  $0 < r < k$ , use  $b$  blocks with frequencies  $0, 1, \dots, v - r - 1, v - r + 1, \dots, v$ . One has sum of frequencies  $bk = \binom{v+1}{2} - (v - r) = \binom{v}{2} + r$ , as required. In each case, the smallest possible maximum frequency  $\mu(1, k, v)$  is realized.

CASE 2:  $2k > v$ . We first show that the given value is a lower bound on  $\mu(1, k, v)$ . Suppose  $m$  is the maximum frequency in an adesign  $A(1, k, v)$ . Then

$$mk \leq bk \leq (m - v + 1) + \dots + (m - 1) + m = mv - \binom{v}{2}.$$

In other words,  $m$  is an integer with  $m(v - k) \geq \binom{v}{2}$ . Conversely, we must realize the given value  $m = \lceil \frac{1}{v-k} \binom{v}{2} \rceil$  as the maximum frequency in an adesign  $A(1, k, v)$ . Put  $bk = mv - \binom{v}{2} - r$ , for some positive integer  $b$  and  $0 \leq r < k$ . Again, use the strategy preceding the statement of the theorem, with  $b = \frac{1}{k}(mv - \binom{v}{2} - r)$  blocks and frequencies

$$m - v, \dots, m - v - r - 1, m - v - r + 1, \dots, m.$$

It remains to check that  $m \leq b$ . However, this follows since  $m$  is the least integer with  $m(v - k) \geq \binom{v}{2}$ . Therefore,  $mk \leq mv - \binom{v}{2}$ . On the other hand,  $b$  is the greatest integer so that  $bk \leq mv - \binom{v}{2}$ .  $\square$

We now turn to adesigns with  $t = 2$ . An important tool is PBD closure. Let  $K$  be a set of positive integers, each at least two. A *pairwise balanced design*  $\text{PBD}(v, K)$  is a set of  $v$  points, together with a set of blocks whose sizes are in  $K$ , having the property that every pair of different points is contained in exactly one block. By ‘breaking up blocks’, it follows that if there exists a  $\text{PBD}(v, K)$  and an  $S_\lambda(2, k, u)$  for every  $u \in K$ , then there exists an  $S_\lambda(2, k, v)$ . This is the foundation of Wilson’s existence theory for block designs; see [5].

It was observed in [1] that adesigns can actually obey a similar recursion. The basic idea is to place adesigns (instead of designs) on the blocks of a PBD. However, each such adesign needs to be accompanied with a block design on those points with sufficiently large  $\lambda$  so as to ‘spread out’ the pairwise frequencies. When restated using  $\mu$ , one obtains the following result.

**Theorem 2.2.** *Suppose there exists a  $\text{PBD}(v, K)$  with  $b$  blocks of sizes  $u_1, u_2, \dots, u_b$ . Put  $M_0 = 0$  and for  $0 < i \leq b$ ,*

$$M_i = \min\{\lambda \geq M_{i-1} : \exists S_\lambda(2, k, u_i)\} + \mu(2, k, u_i).$$

*Then  $\mu(2, k, v) \leq M_b$ .*

We digress here for a few remarks. First, for any  $k$  and sufficiently large  $v$ , there exists  $\text{PBD}(v, \{k+1, k+2, k+3\})$ . This follows easily from Wilson’s theory [5] since  $\gcd\{k(k+1), (k+1)(k+2), (k+2)(k+3)\} = 2$  and  $\gcd\{k, k+1, k+2\} = 1$ . Second, we may let  $\lambda_{\min}(v, k)$  denote the smallest positive integer  $\lambda$  such that there exists an  $S_\lambda(2, k, v)$ . This is well-defined since  $\lambda = \binom{v-2}{k-2}$  always affords a 2-design with  $v$  points and block size  $k$ . Finally, the number of blocks  $b$  of a  $\text{PBD}(v, K)$  is always of order  $v^2$ . In fact,  $b \leq \binom{v}{2} / \binom{u}{2}$ , where  $u$  is the minimum integer in  $K$ . We may now state and prove our main result on 2-adesigns.

**Corollary 2.3.** *There is a constant  $C = C(k)$  such that  $\mu(2, k, v) \leq C \binom{v}{2}$ .*

*Proof.* For  $v$  sufficiently large, apply Theorem 2.2 to a  $\text{PBD}(v, \{k+1, k+2, k+3\})$ , using  $A(2, k, k+j)$ ,  $j = 1, 2, 3$ . Let  $m = \max\{\mu(2, k, k+j) : j = 1, 2, 3\}$  and  $l = \max\{\lambda_{\min}(k+j, k) : j = 1, 2, 3\}$ . Observe  $l$  and  $m$  depend only on  $k$ . Then  $\mu(2, k, v) \leq (l-1)mb \leq C(k) \binom{v}{2}$ .  $\square$

It should be stressed that in practice the best possible upper bound on  $\mu(2, k, v)$  will not be obtained from this exact technique. For instance,  $\mu(2, 3, v)$  was completely determined in [2] using a blend of PBD closure,

group divisible designs, and a variation on “anti-magic cubes”. As expected, one has for all  $v > 3$ ,

$$\mu(2, 3, v) = \begin{cases} \binom{v}{2}, & \text{if } v = 4 \text{ or } v \equiv 2 \pmod{3}, \\ \binom{v}{2} - 1, & \text{otherwise.} \end{cases}$$

We omit further analysis of  $\mu(2, k, v)$  until general constructions surface for  $v$  small relative to  $k$ . Given  $k$  and good enough examples of  $A(2, k, u)$  for various  $u$ , the complete determination of  $\mu(2, k, v)$  could be reduced to a finite problem. The conjectured behavior is  $\mu(2, k, v) = \binom{v}{2} - 1$  for fixed  $k$  and sufficiently large  $v$ .

### 3 General $t$

The theory of pairwise balanced designs does not nicely extend beyond  $t = 2$ . In fact, general  $t$ -designs seem rare unless  $\lambda$  is large. Nonetheless, we begin this section by observing that set systems which are “close” to  $t$ -designs can be combined to yield  $t$ -adesigns.

**Lemma 3.1.** *Suppose there exist integers  $\lambda_1 > \lambda_2$  and a set system  $(V, \widehat{\mathcal{B}})$  such that*

- *the blocks in  $\widehat{\mathcal{B}}$  all have size  $k$ ;*
- *a certain  $t$ -subset  $T \subset V$  is contained in exactly  $\lambda_1$  blocks; and*
- *all other  $t$ -subsets are contained in exactly  $\lambda_2$  blocks.*

*Then*

$$\mu(t, k, v) \leq \left[ \binom{v}{t} - 1 \right] \left( \lambda_1 + \frac{1}{2} \lambda_2 \left[ \binom{v}{t} - 2 \right] \right).$$

*Proof.* List the  $t$ -subsets of  $V$  as  $T_0, \dots, T_{\binom{v}{t}-1}$ . Take copies of the hypothesized set system for  $T = T_i$  with each block having multiplicity  $i$ , where  $i = 0, 1, \dots, \binom{v}{t} - 1$ . Then the number of blocks in total which contain  $T_i$  is

$$f(T_i) = (\lambda_1 - \lambda_2)i + \sum_{j=0}^{\binom{v}{t}-1} \lambda_2 j = (\lambda_1 - \lambda_2)i + \frac{1}{2} \lambda_2 \binom{v}{t} \left[ \binom{v}{t} - 1 \right].$$

Each  $f(T_i)$  is distinct and bounded above by  $f(T_{\binom{v}{t}-1})$ , which simplifies to the asserted bound on  $\mu(t, k, v)$ .  $\square$

Examples of these set systems which favor exactly one  $t$ -subset can be obtained through some elementary linear algebra.

**Theorem 3.2.** *For any  $v > k > t$ , there exists a set system as in Lemma 3.1 with  $\lambda_2 < \lambda_1 < C(t, k)v^{kt+2k-t}$ .*

*Proof.* Fix a  $t$ -subset  $T \subset V$  and consider the family  $\mathcal{B}_j$  of all  $k$ -subsets  $K$  with  $|K \cap T| = t - j$ . Consider another  $t$ -subset  $S \subset V$  with  $|S \cap T| = t - i$ . The number of sets  $K \in \mathcal{B}_j$  which contain  $S$  is a simple count. There are  $\binom{i}{j}$  extensions of  $S \cap T$  to  $K \cap T$  and  $\binom{v-t-i}{v-k-j}$  extensions of  $S \setminus T$  to  $K \setminus T$ . Therefore, the frequency of  $S$  in  $\mathcal{B}_j$  is

$$a_{ij} = \binom{i}{j} \binom{v-t-i}{v-k-j}.$$

Now, let  $A$  denote the matrix  $[a_{ij}]$ , where rows and columns are indexed by  $0, 1, \dots, t$ . Observe that the row sum of  $A$  is  $\binom{v-t}{k-t}$ , a constant, by the Vandermonde convolution. Now define

$$\xi_j = (-1)^j \frac{k-t}{k-t+j} \binom{v-t}{v-k-j}^{-1}.$$

We claim that  $\xi = [\xi_j]_{j=0,1,\dots,t}$  is the first column of  $A^{-1}$ . First, if  $i = 0$ ,  $\sum_{j=0}^t a_{0j} \xi_j = a_{00} \xi_0 = 1$ . On the other hand, if  $i > 0$ , it follows after some binomial identities that

$$\begin{aligned} \sum_{j=0}^t a_{ij} \xi_j &= \sum_{j=0}^t (-1)^j \frac{k-t}{k-t+j} \binom{i}{j} \binom{v-t-i}{v-k-j} \binom{v-t}{v-k-j}^{-1} \\ &= (k-t) \binom{v-t}{i}^{-1} \sum_{j=0}^t (-1)^j \frac{1}{k-t+j} \binom{i}{j} \binom{k-t+j}{i} \\ &= \frac{k-t}{i} \binom{v-t}{i}^{-1} \sum_{j=0}^i (-1)^j \binom{i}{i-j} \binom{k-t+i-1-(i-j)}{i-1} \\ &= 0. \end{aligned}$$

Let  $\mathbf{1}$  denote the all ones vector of dimension  $t+1$ . We may pick constants  $\alpha, \beta$  such that  $\alpha \xi + \beta \mathbf{1}$  is nonnegative and integral. In more detail, choosing  $\alpha = \text{lcm}\{(k-t+j) \binom{v-t}{k-t+j} : j = 0, 1, \dots, t\} = O(v^{k(t+1)})$  ensures that  $\alpha \xi$  is integral, and choosing also a suitable  $\beta = O(v^{k(t+1)})$  ensures that  $\alpha \xi + \beta \mathbf{1}$  is nonnegative. Now form  $\widehat{\mathcal{B}}$  by taking each family  $\mathcal{B}_j$  above with multiplicity  $\alpha \xi_j + \beta$ . The  $t$ -subset frequencies for  $\widehat{\mathcal{B}}$  are, by construction,

$$f(T) = \alpha + \beta \binom{v-t}{k-t},$$

and

$$f(S) = \beta \binom{v-t}{k-t}, \quad S \neq T.$$

Each of these frequencies is at most  $Cv^{k(t+1)}v^{k-t} = Cv^{kt+2k-t}$ , where  $C = C(k, t)$ .  $\square$

As a result, we obtain a crude polynomial upper bound (in  $v$ ) on the maximum frequency in adesigns.

**Corollary 3.3.** *For fixed  $t$  and  $k$ ,  $\mu(t, k, v) < O(v^{(k+1)(t+2)})$ .*

*Proof.* Substitute the frequencies  $\lambda_1, \lambda_2 \leq O(v^{kt+2k-t})$  obtained from Theorem 3.2 into Lemma 3.1. Then  $\mu(t, k, v) \leq O(v^{kt+2k+t}) < O(v^{(k+1)(t+2)})$ .  $\square$

Once again, we do not feel that substantial improvements are forthcoming without a much deeper analysis. It is unfortunate that the linear algebraic solution above seems to beat any bounds we found coming from combinatorial construction techniques. However, this is perhaps not unexpected due to the lack of an existence theory for  $t$ -designs with  $t > 2$ .

## References

- [1] P.J. Dukes, PBD closure for adesigns and asymptotic existence of Sarvate-Beam triple systems. *Bull. Inst. Combin. Appl.* **54** (2008), 5–10.
- [2] P.J. Dukes and J.A. Short-Gershman, A complete existence theory for Sarvate-Beam triple systems. *Australas. J. Combin.* **54** (2012), 261–272.
- [3] K. O’Bryant, A complete annotated bibliography of work related to Sidon sequences. *Elec. J. Combin.* #DS11 (2004), 39pp.
- [4] D. Sarvate and W. Beam, A new type of block design. *Bull. Inst. Combin. Appl.* **50** (2007), 26–28.
- [5] R.M. Wilson, An existence theory for pairwise balanced designs: II, The structure of PBD-closed sets and the existence conjectures. *J. Comb. Theory, Ser. A* **13** (1972), 246–273.